Exhibit 3

Introductory Econometrics

A Modern Approach

4e Jeffrey M. Wooldridge

Introductory Econometrics

A Modern Approach

4e

Jeffrey M. Wooldridge
Michigan State University

SOUTH-WESTERN
CENGAGE Learning™

Australia • Brazil • Japan • Korea • Mexico • Singapore • Spain • United Kingdom • United States

Brief Contents

CHAPTER 1

The Nature of Econometrics and Economic Data

Chapter 1 discusses the scope of econometrics and raises general issues that arise in the application of econometric methods. Section 1.3 examines the kinds of data sets that are used in business, economics, and other social sciences. Section 1.4 provides an intuitive discussion of the difficulties associated with the inference of causality in the social sciences.

1.1 What Is Econometrics?

Imagine that you are hired by your state government to evaluate the effectiveness of a publicly funded job training program. Suppose this program teaches workers various ways to use computers in the manufacturing process. The twenty-week program offers courses during nonworking hours. Any hourly manufacturing worker may participate, and enrollment in all or part of the program is voluntary. You are to determine what, if any, effect the training program has on each worker's subsequent hourly wage.

Now, suppose you work for an investment bank. You are to study the returns on different investment strategies involving short-term U.S. treasury bills to decide whether they comply with implied economic theories.

The task of answering such questions may seem daunting at first. At this point, you may only have a vague idea of the kind of data you would need to collect. By the end of this introductory econometrics course, you should know how to use econometric methods to formally evaluate a job training program or to test a simple economic theory.

Econometrics is based upon the development of statistical methods for estimating economic relationships, testing economic theories, and evaluating and implementing government and business policy. The most common application of econometrics is the forecasting of such important macroeconomic variables as interest rates, inflation rates, and gross domestic product. Whereas forecasts of economic indicators are highly visible and often widely published, econometric methods can be used in economic areas that have nothing to do with macroeconomic forecasting. For example, we will study the effects of political campaign expenditures on voting outcomes. We will consider the effect of school spending on student performance in the field of education. In addition, we will learn how to use econometric methods for forecasting economic time series.

We now demonstrate that the availability of an instrumental variable can be used to consistently estimate the parameters in equation (15.2). In particular, we show that assumptions (15.4) and (15.5) serve to *identify* the parameter β_1. **Identification** of a parameter in this context means that we can write β_1 in terms of population moments that can be estimated using a sample of data. To write β_1 in terms of population covariances, we use equation (15.2): the covariance between z and y is

$$\mathrm{Cov}(z,y) = \beta_1 \mathrm{Cov}(z,x) + \mathrm{Cov}(z,u).$$

Now, under assumption (15.4), $\mathrm{Cov}(z,u) = 0$, and under assumption (15.5), $\mathrm{Cov}(z,x) \neq 0$. Thus, we can solve for β_1 as

$$\beta_1 = \frac{\mathrm{Cov}(z,y)}{\mathrm{Cov}(z,x)}. \tag{15.9}$$

[Notice how this simple algebra fails if z and x are uncorrelated, that is, if $\mathrm{Cov}(z,x) = 0$.] Equation (15.9) shows that β_1 is the population covariance between z and y divided by the population covariance between z and x, which shows that β_1 is identified. Given a random sample, we estimate the population quantities by the sample analogs. After canceling the sample sizes in the numerator and denominator, we get the **instrumental variables (IV) estimator** of β_1:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (z_i - \bar{z})(y_i - \bar{y})}{\sum_{i=1}^{n} (z_i - \bar{z})(x_i - \bar{x})}. \tag{15.10}$$

Given a sample of data on x, y, and z, it is simple to obtain the IV estimator in (15.10). The IV estimator of β_0 is simply $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$, which looks just like the OLS intercept estimator except that the slope estimator, $\hat{\beta}_1$, is now the IV estimator.

It is no accident that when $z = x$ we obtain the OLS estimator of β_1. In other words, when x is exogenous, it can be used as its own IV, and the IV estimator is then identical to the OLS estimator.

A simple application of the law of large numbers shows that the IV estimator is consistent for β_1: $\mathrm{plim}(\hat{\beta}_1) = \beta_1$, provided assumptions (15.4) and (15.5) are satisfied. If either assumption fails, the IV estimators are not consistent (more on this later). One feature of the IV estimator is that, when x and u are in fact correlated—so that instrumental variables estimation is actually needed—it is essentially never unbiased. This means that, in small samples, the IV estimator can have a substantial bias, which is one reason why large samples are preferred.

Statistical Inference with the IV Estimator

Given the similar structure of the IV and OLS estimators, it is not surprising that the IV estimator has an approximate normal distribution in large sample sizes. To perform inference on β_1, we need a standard error that can be used to compute t statistics and confidence intervals. The usual approach is to impose a homoskedasticity assumption, just as in the case of OLS. Now, the homoskedasticity assumption is stated conditional on the instrumental variable, z, not the endogenous explanatory variable, x. Along with the previous assumptions on u, x, and z, we add

$$\mathrm{E}(u^2|z) = \sigma^2 = \mathrm{Var}(u). \tag{15.11}$$

Before we consider a general two-equation SEM, it is useful to gain intuition by considering a simple supply and demand example. Write the system in equilibrium form (that is, with $q_s = q_d = q$ imposed) as

$$q = \alpha_1 p + \beta_1 z_1 + u_1 \tag{16.15}$$

and

$$q = \alpha_2 p + u_2. \tag{16.16}$$

For concreteness, let q be per capita milk consumption at the county level, let p be the average price per gallon of milk in the county, and let z_1 be the price of cattle feed, which we assume is exogenous to the supply and demand equations for milk. This means that (16.15) must be the supply function, as the price of cattle feed would shift supply ($\beta_1 < 0$) but not demand. The demand function contains no observed demand shifters.

Given a random sample on (q, p, z_1), which of these equations can be estimated? That is, which is an **identified equation**? It turns out that the *demand* equation, (16.16), is identified, but the supply equation is not. This is easy to see by using our rules for IV estimation from Chapter 15: we can use z_1 as an IV for price in equation (16.16). However, because z_1 appears in equation (16.15), we have no IV for price in the supply equation.

Intuitively, the fact that the demand equation is identified follows because we have an observed variable, z_1, that shifts the supply equation while not affecting the demand equation. Given variation in z_1 and no errors, we can trace out the demand curve, as shown in Figure 16.1. The presence of the unobserved demand shifter u_2 causes us to estimate the demand equation with error, but the estimators will be consistent, provided z_1 is uncorrelated with u_2.

FIGURE 16.1

Shifting supply equations trace out the demand equation. Each supply equation is drawn for a different value of the exogenous variable, z_1.